

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-229740

(43)Date of publication of application : 16.08.2002

(51)Int.Cl.

G06F 3/06

G06F 12/16

G06F 13/00

(21)Application number : 2001-023974

(71)Applicant : TOSHIBA CORP

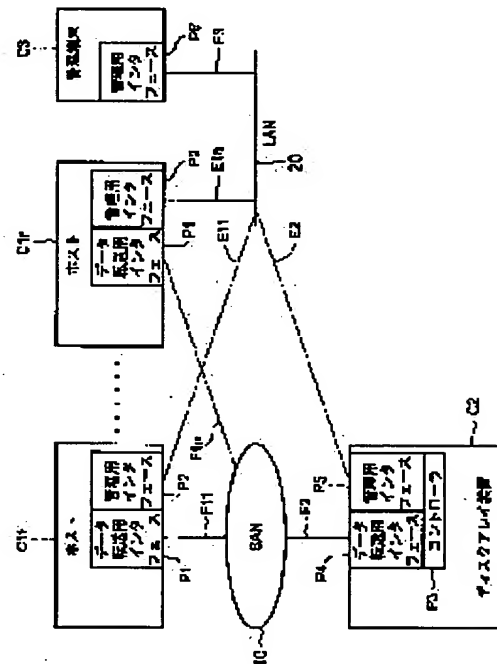
(22)Date of filing : 31.01.2001

(72)Inventor : YAMASHITA AKIRA

(54) STORAGE SYSTEM WITH HIGH AVAILABILITY CONTROL INTERFACE, HOST COMPUTER USED THEREFOR, AND COMPUTER STORED WITH DRIVER PROGRAM EXECUTED BY COMPUTER

(57)Abstract:

PROBLEM TO BE SOLVED: To reduce the number of components necessary for multiplexing of data transfer path by sharing a control path for a data transfer path.  
SOLUTION: A duplex driver in a host computer C11, controls in data transferring between the host computer C11 and a disk array device C2, by switching a path F led to a data transfer interface P4 of a disk array device C2 via a cable F11, SAN 10 and a cable F2 from a data transfer interface P1 of the host computer C11 and a path E from a control interface P2 of the host computer C11 led to a control interface P5 of the disk array device C2 via a cable E11, LAN 20 and a cable E2, as a data transfer path.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's

decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-229740

(P2002-229740A)

(43) 公開日 平成14年 8月16日 (2002.8.16)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	タームコード* (参考)
G 0 6 F 3/06	3 0 1	G 0 6 F 3/06	3 0 1 B 5 B 0 1 8
12/16	3 2 0	12/16	3 2 0 L 5 B 0 6 5
13/00	3 0 1	13/00	3 0 1 P 5 B 0 8 3

審査請求 未請求 請求項の数10 O L (全 11 頁)

(21) 出願番号 特願2001-23974(P2001-23974)

(22) 出願日 平成13年 1月31日 (2001.1.31)

(71) 出願人 000003078

株式会社東芝

東京都港区芝浦一丁目 1 番 1 号

(72) 発明者 山下 亮

東京都府中市東芝町 1 番地 株式会社東芝

府中事業所内

(74) 代理人 100058479

弁理士 鈴江 武彦 (外 6 名)

Fターム(参考) 5B018 GA06 HA35 HA40 MA14

5B065 BA01 CA05 CA30

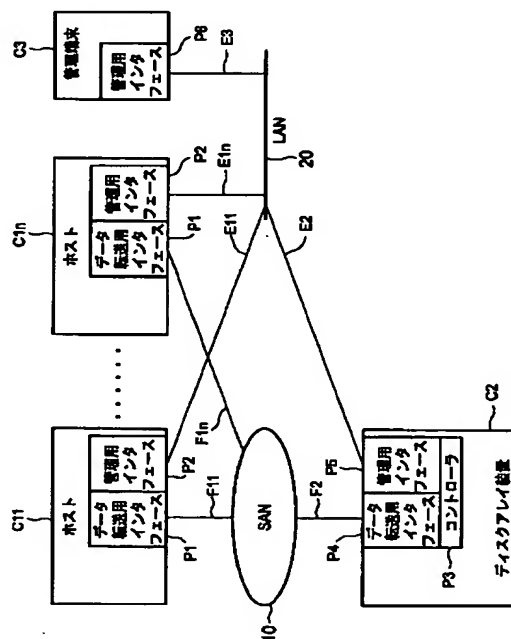
5B083 AA08 BB03 CC04 CD11 EE08

(54) 【発明の名称】 管理用インタフェースを用いた高可用性ストレージシステム、同システムに用いられるホストコンピュータ、及び同コンピュータにより実行されるドライバプログラムを記憶したコンピュータ

# (57) 【要約】

【課題】 管理用バスをデータ転送用バスに共用することにより、データ転送用バスの多重化に必要な構成要素の削減を図る。

【解決手段】 ホストC11内の二重化ドライバは、ホストC11とディスクアレイ装置C2との間のデータ転送で、そのデータ転送用のバスとして、ホストC11のデータ転送用インタフェースP1からケーブルF11、SAN10及びケーブルF2を経由してディスクアレイ装置C2のデータ転送用インタフェースP4に至るバスFと、ホストC11の管理用インタフェースP2からケーブルE11、LAN20及びケーブルE2を経由してディスクアレイ装置C2の管理用インタフェースP5に至るバスEとを切り替え制御する。



【特許請求の範囲】

【請求項 1】 データ転送のための第 1 のインタフェースと、状態監視を含むシステム管理のための第 2 のインタフェースとを有する少なくとも 1 つのホストコンピュータと、

データ転送のための第 3 のインタフェースと、状態監視を含むシステム管理のための第 4 のインタフェースとを有する少なくとも 1 つのストレージ装置と、

前記ホストコンピュータの前記第 1 のインタフェースと前記ストレージ装置の前記第 3 のインタフェースとを接続するデータ転送用接続インタフェースと、

前記ホストコンピュータの前記第 2 のインタフェースと前記ストレージ装置の前記第 4 のインタフェースとを接続する管理用接続インタフェースとを具備し、

前記ホストコンピュータは、当該コンピュータと前記ストレージ装置との間のデータ転送で、そのデータ転送用のバスとして、当該コンピュータの前記第 1 のインタフェースから前記データ転送用接続インタフェースを経由して前記ストレージ装置の前記第 3 のインタフェースに至る第 1 のバスと、当該コンピュータの前記第 2 のインタフェースから前記管理用接続インタフェースを経由して前記ストレージ装置の前記第 4 のインタフェースに至る第 2 のバスとを切り替え制御する二重化ドライバを有することを特徴とする高可用性ストレージシステム。

【請求項 2】 前記ホストコンピュータは、前記第 2 のインタフェース及び前記管理用接続インタフェースを利用しての管理データ転送と、前記管理用接続インタフェースを含む前記第 2 のバスを利用しての前記ストレージ装置との間のデータ転送とを制御する管理用ドライバを有することを特徴とする請求項 1 記載の高可用性ストレージシステム。

【請求項 3】 前記ストレージ装置は、前記第 4 のインタフェース及び前記管理用接続インタフェースを利用しての管理データ転送と、前記管理用接続インタフェースを含む前記第 2 のバスを利用しての前記ホストコンピュータとの間のデータ転送とを制御する管理用ドライバを有することを特徴とする請求項 1 記載の高可用性ストレージシステム。

【請求項 4】 前記ホストコンピュータの管理用ドライバは、前記第 2 のバスを利用してのデータ転送では、前記ストレージ装置に転送すべき前記データ転送用接続インタフェースでのデータ転送に適用されるデータ転送用プロトコルのディスクアクセスコマンド及びディスクアクセスデータを、前記管理用接続インタフェースを介して前記ストレージ装置に転送するための特定データ形式に変換すると共に、前記第 2 のインタフェースにより前記データ転送用接続インタフェースから受信された前記特定データ形式のデータを前記データ転送用プロトコルのデータに変換することを特徴とする請求項 2 記載の高可用性ストレージシステム。

【請求項 5】 前記ストレージ装置の管理用ドライバ

は、前記第 2 のバスを利用してのデータ転送では、前記ホストコンピュータに転送すべき前記データ転送用接続インタフェースでのデータ転送に適用されるデータ転送用プロトコルのディスクアクセスコマンド及びディスクアクセスデータを、前記管理用接続インタフェースを介して前記ホストコンピュータに転送するための特定データ形式に変換すると共に、前記第 4 のインタフェースにより前記データ転送用接続インタフェースから受信された前記特定データ形式のデータを前記データ転送用プロトコルのデータに変換することを特徴とする請求項 3 記載の高可用性ストレージシステム。

【請求項 6】 データ転送のための第 1 のインタフェースと、状態監視を含むシステム管理のための第 2 のインタフェースとを有し、データ転送のための第 3 のインタフェースと、状態監視を含むシステム管理のための第 4 のインタフェースとを有するストレージ装置と、前記第 1 のインタフェース、データ転送用接続インタフェース、及び前記第 3 のインタフェースにより接続されると共に、前記第 2 のインタフェース、管理用接続インタフェース、及び前記第 4 のインタフェースにより接続される、高可用性ストレージシステムに用いられるホストコンピュータであって、前記コンピュータと前記ストレージ装置との間のデータ転送で、そのデータ転送用のバスとして、当該コンピュータの前記第 1 のインタフェースから前記データ転送用接続インタフェースを経由して前記ストレージ装置の前記第 3 のインタフェースに至る第 1 のバスと、当該コンピュータの前記第 2 のインタフェースから前記管理用接続インタフェースを経由して前記ストレージ装置の前記第 4 のインタフェースに至る第 2 のバスとを切り替え制御する二重化ドライバを具備することを特徴とする高可用性ストレージシステムに用いられるホストコンピュータ。

【請求項 7】 前記第 2 のインタフェース及び前記管理用接続インタフェースを利用しての管理データ転送と、前記管理用接続インタフェースを含む前記第 2 のバスを利用しての前記ストレージ装置との間のデータ転送とを制御する管理用ドライバを更に具備することを特徴とする請求項 6 記載の高可用性ストレージシステムに用いられるホストコンピュータ。

【請求項 8】 データ転送のための第 1 のインタフェースと、状態監視を含むシステム管理のための第 2 のインタフェースとを有し、データ転送のための第 3 のインタフェースと、状態監視を含むシステム管理のための第 4 のインタフェースとを有するストレージ装置と、前記第 1 のインタフェース、データ転送用接続インタフェース、及び前記第 3 のインタフェースにより接続されると共に、前記第 2 のインタフェース、管理用接続インタフェース、及び前記第 4 のインタフェースにより接続される、高可用性ストレージシステムに用いられるホストコ

ンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体であって、前記コンピュータを、  
前記コンピュータの前記第1のインタフェースから前記データ転送用接続インタフェースを経由して前記ストレージ装置の前記第3のインタフェースに至る第1のバスと、前記コンピュータの前記第2のインタフェースから前記管理用接続インタフェースを経由して前記ストレージ装置の前記第4のインタフェースに至る第2のバスとを監視する監視手段と、  
前記コンピュータと前記ストレージ装置との間のデータ転送で、そのデータ転送用のバスとして、前記第1のバスと前記第2のバスとを前記監視ステップの監視結果に基づいて切り替え制御する切り替え制御手段として機能させるためのドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【請求項9】 前記コンピュータを、  
前記第2のインタフェース及び前記管理用接続インタフェースを利用した管理データ転送と、前記管理用接続インタフェースを含む前記第2のバスを利用した前記ストレージ装置との間のデータ転送とを制御する転送制御手段として機能させるための請求項8記載のドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【請求項10】 データ転送のための第1のインタフェースと、状態監視を含むシステム管理のための第2のインタフェースとを有し、データ転送のための第3のインタフェースと、状態監視を含むシステム管理のための第4のインタフェースとを有するストレージ装置と、前記第1のインタフェース、データ転送用接続インタフェース、及び前記第3のインタフェースにより接続されると共に、前記第2のインタフェース、管理用接続インタフェース、及び前記第4のインタフェースにより接続される、高可用性ストレージシステムに用いられるホストコンピュータにおける二重化バス制御方法であって、  
前記コンピュータの前記第1のインタフェースから前記データ転送用接続インタフェースを経由して前記ストレージ装置の前記第3のインタフェースに至る第1のバスと、前記コンピュータの前記第2のインタフェースから前記管理用接続インタフェースを経由して前記ストレージ装置の前記第4のインタフェースに至る第2のバスとを監視するステップと、  
前記コンピュータと前記ストレージ装置との間のデータ転送で、そのデータ転送用のバスとして、前記第1のバスと前記第2のバスとを前記監視ステップの監視結果に基づいて切り替え制御するステップとを具備することを特徴とする二重化バス制御方法

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、ストレージ装置ア

クセスのためのデータ転送用のバスと、状態監視を含むシステム管理用のバスとを独立して有する高可用性ストレージシステム、同システムに用いられるホストコンピュータ、及び同コンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体に関する。

【0002】

【従来の技術】 一般に、ストレージ装置、例えば複数のディスク装置を備えたディスクアレイ装置は、ホストコンピュータ（以下、ホストと略称する）からのデータアクセスに用いられるインタフェース（データ転送用インタフェース）として、SCSI（Small Computer System Interface）、或いはファイバチャネルなどのインタフェースを少なくとも1つ実装している。

【0003】 ディスクアレイ装置を用いて構成されるシステム（ストレージシステム）には、高可用性を実現するために、ホストとディスクアレイ装置との間を複数のバス、例えば2つのバス（データ転送用バス）で接続しているものがある。このようなシステムは高可用性ストレージシステムと呼ばれ、ホストとディスクアレイ装置のそれぞれに、バス数分のデータ転送用インタフェースが実装される。この高可用性ストレージシステムでは、稼働中のバス（実行バス）にケーブルの断線などの何らかの障害が発生した場合に、別のバス（待機バス）を使用して運用を継続させることができるように、複数のバスの運用をホストにて制御することによって、システムの可用性を向上させている。

【0004】

【発明が解決しようとする課題】 上記したように、従来のディスクアレイ装置を備えた高可用性ストレージシステムでは、高可用性を実現するために、ホストとディスクアレイ装置にそれぞれデータ転送用インタフェースを複数、例えば2つ実装し、ホストとディスクアレイ装置との間を2つのデータ転送用バスで接続して二重化バス構成とすることで、データ転送用バスの障害に対応できるようにしている。

【0005】 このようなストレージシステムでは、2つのデータ転送用バスの一方に障害が発生しても、もう一方のバスを使用することで、確かにシステムの可用性を向上することができる。しかし、ホストとディスクアレイ装置にそれぞれデータ転送用インタフェースを2つ実装して二重化バス構成とすることは、ハードウェア構成要素数の増加を招くという問題がある。

【0006】 一方、高可用性ストレージシステムに用いられるディスクアレイ装置には、電源、冷却用ファンなどの構成要素の状態や、データのアクセス状況などの状態を外部の管理端末或いは管理用ホストに通知するための、例えばCSMA/CD（Carrier Sense Multiple Access with Collision Detection）方式を適用する管理用インタフェースを実装しているものがある。特に近時

は、システムの管理のために、ディスクアレイ装置だけでなく、ホストを含めて、全てのシステム構成機器の状態監視が行えるように、各システム構成機器に管理用インタフェースを実装してLAN（ローカルエリアネットワーク）と接続することにより管理用バスを設定し、各システム構成機器の状態監視を当該管理用バスを介して1つの管理端末或いは1つの管理用ホスト（ネットワーク接続機器）にて集中して行える高可用性ストレージシステムが実用化されている。

【0007】本発明は上記事情を考慮してなされたものでその目的は、ストレージ装置を含むシステム構成機器の状態監視等に用いられる管理用バスをホストとストレージ装置との間のデータ転送用バスに共用することにより、データ転送用バスの多重化に必要な構成要素を削減でき、構成要素を削減しない場合には多重化数を1つ増加してより一層可用性が向上できる高可用性ストレージシステム、同システムに用いられるホストコンピュータ、及び同コンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体を提供することにある。

【0008】

【課題を解決するための手段】本発明に係る高可用性ストレージシステムは、データ転送のための第1のインタフェースと、状態監視を含むシステム管理のための第2のインタフェースと、二重化ドライバとを有する少なくとも1つのホストコンピュータと、データ転送のための第3のインタフェースと、状態監視を含むシステム管理のための第4のインタフェースとを有する少なくとも1つのストレージ装置と、上記ホストコンピュータの第1のインタフェースと上記ストレージ装置の第3のインタフェースとを接続するデータ転送用接続インタフェースと、上記ホストコンピュータの第2のインタフェースと上記ストレージ装置の第4のインタフェースとを接続する管理用接続インタフェースとを備え、上記コンピュータの二重化ドライバが、当該コンピュータと上記ストレージ装置との間のデータ転送で、そのデータ転送用のバスとして、当該コンピュータの第1のインタフェースーデータ転送用接続インタフェースーストレージ装置の第3のインタフェースの経路である第1のバスと、当該コンピュータの第2のインタフェースーデータ転送用接続インタフェースーストレージ装置の第4のインタフェースの経路である第2のバスとを切り替え制御することを特徴とする。

【0009】このような構成においては、データ転送用の上記第1のバスに加えて、従来であれば管理データの転送のみに用いられていた管理用のバスのうちの上記第2のバスを、第1のバスと同様のデータ転送用バスとしても割り当てることができる。この結果、従来とは異なって、二重化のためのハードウェア構成、即ちホストコンピュータ及びストレージ装置内に、第1及び第3のイ

ンタフェースに相当する別のデータ転送用インタフェースを設ける必要がなく、ハードウェア構成の削減が図れ、しかも従来の構成と同等のシステムの可用性を実現できる。ここで、ホストコンピュータ及びストレージ装置内に、それぞれ上記第1及び第3のインタフェースを複数設けて、上記第1のバス自体を多重化するならば、より一層システムの可用性を高めることができる。

【0010】また、上記二重化ドライバによる制御のためには、ホストコンピュータ内に上記第1及び第2のバスを監視する手段を設けるとよい。

【0011】また、上記コンピュータ内に、第2のインタフェース及び管理用接続インタフェースを利用しての管理データ転送と、管理用接続インタフェースを含む第2のバスを利用してのストレージ装置との間のデータ転送とを制御する管理用ドライバを設けるとよい。同様に、上記ストレージ装置の例えばコントローラ内に、第4のインタフェース及び管理用接続インタフェースを利用しての管理データ転送と、管理用接続インタフェースを含む第2のバスを利用してのホストコンピュータとの間のデータ転送とを制御する管理用ドライバを設けるとよい。

【0012】以上の構成の高可用性ストレージシステムに係る発明は、当該システム中のホストコンピュータに係る発明としても成立する。また、このホストコンピュータに係る発明は、当該コンピュータにより実行されるドライバプログラムに係る発明、及び当該コンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体に係る発明、更には当該コンピュータにおける二重化バス制御方法に係る発明としても成立する。

【0013】

【発明の実施の形態】以下、本発明の実施の形態につき図面を参照して説明する。

【0014】図1は本発明の一実施形態に係る高可用性ストレージシステムの構成を示すブロック図である。図1のシステムは、主として、n台のホスト（ホストコンピュータ）C11～C1nと、ストレージ装置としてのディスクアレイ装置C2と、管理端末C3とから構成される。

【0015】ホストC11～C1n及びディスクアレイ装置C2は、SAN（Storage Area Network：ストレージエリアネットワーク）10により相互接続されている。これによりホストC1～C1nは、SAN10を介してディスクアレイ装置C2を共有することができる。なお、図1では、ディスクアレイ装置C2は1台であるが、複数のディスクアレイ装置C2をSAN10を介してホストC11～C1nと接続することも可能である。

【0016】また、ホストC11～C1n及びディスクアレイ装置C2は、LAN（Local Area Network：ローカルエリアネットワーク）20により管理端末C3と接

続されている。

【0017】ホストC1x (x=1~n)は、データ転送用インタフェースP1及び管理用インタフェースP2の両ハードウェア構成要素を有している。データ転送用インタフェースP1は、ホストC1xとディスクアレイ装置C2とをSAN10にて接続するためインタフェースを提供する。管理用インタフェースP2は、ホストC1xと管理端末C3とをLAN20にて接続するためのインタフェースを提供する。

【0018】ディスクアレイ装置C2は、複数のディスク装置(図示せず)と、これらのディスク装置のアクセス制御及びディスクアレイ装置C2全体の制御を行うコントローラ(ディスクアレイコントローラ)P3とから構成される。コントローラP3は、データ転送用インタフェースP4及び管理用インタフェースP5の両ハードウェア構成要素を有している。データ転送用インタフェースP4は、ディスクアレイ装置C2とホストC1xとをSAN10にて接続するためインタフェースを提供する。管理用インタフェースP5は、ディスクアレイ装置C2と管理端末C3とをLAN20にて接続するためのインタフェースを提供する。

【0019】管理端末C3は、ホストC11~C1n及びディスクアレイ装置C2の状態の監視を含むシステム管理を集中して実行する管理専用の端末である。管理端末C3は、当該端末C3のハードウェア構成要素としての管理用インタフェースP6を有している。管理用インタフェースP6は、管理端末C3とホストC11~C1n及びディスクアレイ装置C2とをLAN20にて接続するためのインタフェースを提供する。

【0020】ホストC11~C1nの各データ転送用インタフェースP1には、それぞれデータ転送用インタフェースケーブルF11~F1nの一端が接続され、ディスクアレイ装置C2のデータ転送用インタフェースP4には、データ転送用インタフェースケーブルF2の一端が接続されている。ケーブルF11~F1n、F2はファイバチャネルケーブル等である。SAN10は、ケーブルF11~F1n、F2の他端をデータ転送用接続インタフェースとしてのファブリックスイッチに接続することにより実現される。つまり、SAN10は実質的にデータ転送用接続インタフェースである。

【0021】ホストC11~C1nの管理用インタフェースP2には、それぞれ管理用インタフェースケーブルE11~E1nの一端が接続され、ディスクアレイ装置C2の管理用インタフェースP5には、管理用インタフェースケーブルE2の一端が接続されている。また、管理端末C3の管理用インタフェースP6には、管理用インタフェースケーブルE3の一端が接続されている。ケーブルE11~E1n、E2、E3はCSMA/CD方式を適用するLANケーブルである。LAN20は、ケーブルE11~E1n、E2、E3の他端を管理用接続

インタフェースとしてのハブに接続することにより実現される。つまり、LAN20は実質的にデータ転送用接続インタフェースである。

【0022】図2は、ホストC1x (x=1~n)の主としてソフトウェア構成を示す。ホストC1xには、アプリケーションソフトウェアS1、オペレーティングシステム(OS)S2、二重化ドライバS3、SCSIドライバS4、S41、データ転送用ドライバS5、及び管理用ドライバS6の各ソフトウェアがインストールされている。ドライバS3~S6からなるドライバ部(ドライバソフトウェア)S7は、例えばCD-ROM等の記憶媒体からホストC1x内の記憶装置に読み込まれたものである。このドライバ部S7が、インターネット等を介してホストC1x内の記憶装置にダウンロードされたものであっても構わない。

【0023】二重化ドライバS3は、SCSIドライバS4、データ転送用ドライバS5及びデータ転送用インタフェースP1から、図1中のケーブルF1x及びF2を経由してディスクアレイ装置C2のデータ転送用インタフェースP4までのバス(以下、バスFと称する)と、SCSIドライバS41、管理用ドライバS6及び管理用インタフェースP2から、ケーブルE1x及びE2を経由してディスクアレイ装置C2の管理用インタフェースP5までのバス(以下、バスEと称する)の切り替えを制御する。このように本実施形態の特徴は、状態監視を含む管理用のバスEが、ホストC1xとディスクアレイ装置C2との間のデータ転送用にも切り替えられる点にある。

【0024】SCSIドライバS4は、SCSIデバイスとして認識されるディスクアレイ装置C2とホストC1xとの間のデータ転送に関する、オペレーティングシステムS2とデータ転送用ドライバS5との間の制御を行う。同様に、SCSIドライバS41は、ディスクアレイ装置C2とホストC1xとの間のデータ転送に関する、オペレーティングシステムS2と管理用ドライバS6との間の制御を行う。なお、データ転送用ドライバS5及び管理用ドライバS6に、それぞれSCSIドライバS4及びS41の機能を持たせることも可能である。

【0025】データ転送用ドライバS5は、ホストC1xのデータ転送用インタフェースP1を制御し、SCSIドライバS4にSCSIデバイスとしてディスクアレイ装置C2を認識させるための制御及びデータの授受を行う。管理用ドライバS6は、ホストC1xの管理用インタフェースP2を制御し、SCSIドライバS41にSCSIデバイスとしてディスクアレイ装置C2を認識させるための制御及びデータの授受を行う。また管理用ドライバS6は、オペレーティングシステムS2に対して、通常のネットワーク機器のインタフェースとして認識させるための制御を行う。

【0026】図3は、ディスクアレイ装置C2内の主と

してコントローラ P 3 のソフトウェア構成（例えばファームウェア構成）を示す。コントローラ P 3 には、オペレーティングシステム S 8、システム制御 S 9、アクセス制御 S 10、状態監視 S 11、データ転送用ドライバ S 12 及び管理用ドライバ S 13 の各ソフトウェアがインストールされている。

【0027】オペレーティングシステム S 8 はディスクアレイ装置 C 2 のコントローラ P 3 に組み込まれた基本ソフトウェアであり、例えば ROM 等の不揮発性メモリに予め格納されている。

【0028】システム制御 S 9 は、ディスクアレイ装置 C 2 全体の制御を行う。アクセス制御 S 10 は、ディスク装置へのアクセスを制御する。状態監視 S 11 は、ディスクアレイ装置 C 2 の電源、ファン、ディスク装置に代表される構成要素の状態、及び運用状況など、ディスクアレイ装置 C 2 の種々の状態を監視する。

【0029】データ転送用ドライバ S 12 は、ディスクアレイ装置 C 2 のデータ転送用インタフェース P 4 とアクセス制御 S 10 との間のデータの授受を制御する。管理用ドライバ S 13 は、ディスクアレイ装置 C 2 の管理用インタフェース P 5 を制御し、状態監視のときは、当該インタフェース P 5 と状態監視 S 10 との間のデータの授受を制御し、ディスクアクセスのときは、当該インタフェース P 5 とアクセス制御 S 9 との間のデータの授受を制御する。

【0030】次に、本実施形態の動作を、（１）ホストの動作と（２）ディスクアレイ装置の動作に分けて順次説明する。

【0031】（１）ホストの動作

まず、ホスト C 1 x（x=1~n）に図 2 中の二重化ドライバ S 3 が実装されていないものとする、そのホスト C 1 x のオペレーティングシステム S 2 からは、ディスクアレイ装置 C 2 がバス F とバス E の 2 つのバスによって 2 台のディスクアレイ装置として認識される。

【0032】これに対して本実施形態のように二重化ドライバ S 3 が実装されたホスト C 1 x では、その二重化ドライバ S 3 により以下の制御が行われる。

【0033】まず二重化ドライバ S 3 は、バス F 及びバス E の各々の経路から認識されるディスクアレイ装置が、同じディスクアレイ装置であることを確認するための処理を行う。この確認は、バス F 及びバス E をそれぞれ介してディスクアレイ装置から固有の情報を読み出すことで行う。ディスクアレイ装置固有の情報とは、ディスクアレイ装置に付けられたシリアル ID 番号、ファイバチャネル規格で定められたファイバチャネル機器の WWN（World Wide Name）など、ディスクアレイ装置を特定することが可能な情報である。なお、SAN 10 を介して複数のディスクアレイ装置 C 2 が接続されている場合には、各ディスクアレイ装置 C 2 毎に固有の情報を読み出せばよい。

【0034】次に二重化ドライバ S 3 は、バス F 及びバス E からそれぞれ認識される同じディスクアレイ装置 C 2 を、仮想の 1 台のディスクアレイ装置（仮想デバイス）D x として、オペレーティングシステム S 2 に認識させる。

【0035】そして、オペレーティングシステム S 2 から仮想デバイス D x に対するディスクアクセスの要求があった場合、二重化ドライバ S 3 は、バス F が正常に動作しているならば、当該バス F を用いてディスクアレイ装置 C 2 へのアクセスを行う。これに対し、ケーブル F 1 x の断線などの要因でバス F に何らかの障害が発生したことの通知を、データ転送用インタフェース P 1、データ転送用ドライバ S 5 及び SCSI ドライバ S 4 を経由して受けたときは、二重化ドライバ S 3 はオペレーティングシステム S 2 に障害の発生を通知し、バス E を用いてディスクアレイ装置 C 2 へのアクセスを行う。

【0036】二重化ドライバ S 3 は、当該ドライバ S 3 の設定情報に設定された時間間隔で定期的にバス F 及びバス E の障害監視を行い、バスの異常を検出した場合には、オペレーティングシステム S 2 に通知を行い、障害の発生していないバスを用いてディスクアクセスを行う。

【0037】また二重化ドライバ S 3 は、バスの障害が回復されたことを検出した場合には、その旨をオペレーティングシステム S 2 に通知する。もし、バス F で発生していた障害が回復した場合には、オペレーティングシステム S 2 はバス E の経路を介して行っていたディスクアレイ装置 C 2 へのディスクアクセスを、バス F の経路でのディスクアクセスに切り替える。

【0038】ホスト C 1 x の管理用ドライバ S 6 は、管理用インタフェース P 2 を制御し、オペレーティングシステム S 2 から制御されるホスト C 1 x の状態を、ケーブル E 1 1 及び E 3 を経由して管理端末 C 3 に通知する。この他に管理用ドライバ S 6 は、本発明に関係する以下の制御を行う。

【0039】即ちホスト C 1 x の管理用ドライバ S 6 は、当該ホスト C 1 x の二重化ドライバ S 3、及び SCSI ドライバ S 4 1 から制御される SCSI プロトコルのディスクアクセスコマンド（例えばリード/ライトコマンド）、ディスクアクセスデータを、LAN 20 を介して転送するための、iSCSI、SCSI over IP などと呼ばれるデータ形式（iSCSI 規格）に変換し、管理用インタフェース P 2 を制御して LAN 20 への転送を行う。また管理用ドライバ S 6 は、LAN 20 から管理用インタフェース P 2 が受信した、ディスクアレイ装置 C 2 の管理用インタフェース P 5 から転送された iSCSI のデータをデータ変換して、SCSI ドライバ S 4 1 に転送する。

【0040】（２）ディスクアレイ装置の動作

ディスクアレイ装置 C 2 内の管理用ドライバ S 13 は、



管理用インタフェース P5 を制御し、状態監視 S11 から制御されるディスクアレイ装置 C2 の状態をケーブル E2 及び E3 を経由して管理端末 C3 に通知する。この他に管理用ドライバ S13 は、本発明に係る以下の制御を行う。

【0041】即ちディスクアレイ装置 C2 の管理用ドライバ S13 は、当該ディスクアレイ装置 C2 のアクセス制御 S10 から制御される SCSI プロトコルのディスクアクセスコマンド（例えば応答コマンド）、ディスクアクセスデータを、LAN20 を介して転送するためのデータ形式（iSCSI 規格）に変換し、管理用インタフェース P5 を制御して LAN20 への転送を行う。また管理用ドライバ S13 は、LAN20 から管理用インタフェース P5 が受信した、ホスト C1x の管理用インタフェース P2 から転送された iSCSI のデータをデータ変換して、アクセス制御 S10 に転送する。

【0042】次に、上記したホスト C1x の二重化ドライバ S3 の動作の詳細を、（a）バスが二重化されたデバイスの認識処理、（b）ディスクアクセス処理、（c）バスの障害監視処理の各々に分けて説明する。

【0043】（a）バスが二重化されたデバイスの認識処理

まず、ホスト C1x の二重化ドライバ S3 による、バスが二重化されたデバイス（ディスクアレイ装置 C2）の認識処理、即ちバス F 及びバス E の各々の経路から認識されるディスクアレイ装置が、同じディスクアレイ装置であることを確認するための処理について、図 4 のフローチャートを参照して説明する。

【0044】今、二重化ドライバ S3 が、アイドル状態（ステップ 100）において、オペレーティングシステム S2 からのディスク接続情報の要求を受信したものとする（ステップ 101）。

【0045】二重化ドライバ S3 は、ディスク接続情報要求に応じて、バス F に接続されたディスクアレイ装置の固有の情報（ディスク識別情報）、及びバス E に接続されたディスクアレイ装置の固有の情報を、それぞれバス F 及びバス E の経路から読み出す。

【0046】次に二重化ドライバ S3 は、バス F から読み出したディスク装置固有の情報と、バス E から読み出したディスク装置固有の情報とを比較して、両情報が一致しているか否かにより、バス F とバス E とが同じディスクアレイ装置に接続されているか否かを判定する（ステップ 103）。

【0047】もし、バス F とバス E とが同じディスクアレイ装置に接続されていないと判定した場合、即ちバス F とバス E とが異なるディスクアレイ装置に接続されていると判定した場合、二重化ドライバ S3 はバスの二重化がされていないものとして、ステップ 104、105 を実行する。二重化ドライバ S3 はこのステップ 104、105 において、バス F のディスクアレイ装置を D

x1 とし、バス E のディスクアレイ装置を Dx2 とする。

【0048】これに対し、バス F とバス E とが同じディスクアレイ装置に接続されていると判定した場合には、即ちバス F に接続されたディスクアレイ装置とバス E に接続されたディスクアレイ装置とが同一のディスクアレイ装置であると判定した場合には、二重化ドライバ S3 は当該ディスクアレイ装置との間でバスの二重化がされているものとして、ステップ 106 を実行する。二重化ドライバ S3 はこのステップ S106 において、バス F に接続され、且つバス E にも接続された同一のディスクアレイ装置（ここではディスクアレイ装置 C11）を仮定のディスクアレイ装置（仮想デバイス）Dx1 とする。

【0049】そして二重化ドライバ S3 は、認識したデバイス（ディスクアレイ装置）の情報 Dx<sub>m</sub>（m=1 または m=2）をオペレーティングシステム S2 に返してアイドル状態（ステップ 100）に戻る。

【0050】（b）ディスクアクセス処理  
次に、二重化ドライバ S3 によるディスクアクセス処理について、図 5 のフローチャート（のステップ 201～209）を参照して説明する。

【0051】まず二重化ドライバ S3 が、アイドル状態（ステップ 100）において、オペレーティングシステム S2 からの Dx1（で示されるディスクアレイ装置）に対するディスクアクセス要求を受信したものとする（ステップ 201）。

【0052】二重化ドライバ S3 は、Dx1 への実行バス（バス F とする）の状態が正常であるか否かを判定する（ステップ 202）。この判定は、後述する障害監視処理（ステップ 301～306）での前回の実行結果として、エラーが通知されているか否かに基づいて行われる。

【0053】二重化ドライバ S3 は、ステップ 202 で実行バスの状態が異常であると判定した場合、Dx1 へのバスが二重化されているか否かを、先のデバイス認識処理（ステップ 101～107）での Dx<sub>y</sub> の設定内容から判定する（ステップ 203）。

【0054】もし、Dx1 へのバス二重化されているならば、二重化ドライバ S3 は待機バス（バス E とする）の状態が正常であるか否かを判定する（ステップ 204）。そして二重化ドライバ S3 は、待機バスが正常であるならば、待機バスと実行バスとを切り替えて、今まで待機バスであったバス（バス E）を実行バスに、実行バスであったバス（バス F）を待機バスに変更する（ステップ 205）。

【0055】これに対し、ステップ 204 で待機バスも異常であると判定した場合には、ディスクアクセス不可能であることから、二重化ドライバ S3 はアクセスが実行できなかったことを示すアクセス異常完了をオペレー

ティングシステム S2 に通知する (ステップ 209)。また、ステップ 203 で D x 1 へのバスが二重化されていないと判定した場合にも、ディスクアクセス不可能であることから、二重化ドライバ S3 はステップ 209 を実行して、アクセス異常完了をオペレーティングシステム S2 に通知する。

【0056】さて、ステップ 202 で実行バスが正常であると判定した場合、或いはステップ 205 で待機バスと実行バスとを切り替えた場合、二重化ドライバ S3 は現在の実行バスを介してオペレーティングシステム S2 から要求されたディスクアクセスを実行する (ステップ 206)。

【0057】二重化ドライバ S3 は、ステップ 206 でのディスクアクセスが正常に実行された場合 (ステップ 207 の YES)、正常にアクセスが完了したことをオペレーティングシステム S2 に通知して (ステップ 208)、アイドル状態 (ステップ 100) に戻る。これに対し、ステップ 206 でのディスクアクセスの実行中に実行バスの異常を検出した場合やその他の要因でディスクアクセスが正常に実行されなかった場合には (ステップ 207 の NO)、二重化ドライバ S3 は待機バスでのディスクアクセスのためにステップ 203 に分岐する。

【0058】(c) バスの障害監視処理  
次に、二重化ドライバ S3 によるバスの障害監視処理について、図 5 のフローチャート (のステップ 301 ~ 306) を参照して説明する。

【0059】二重化ドライバ S3 は、アイドル状態 (ステップ 100) において一定時間が経過する毎に。バスの障害を監視する処理 (ステップ 301 ~ 306) を実行する。即ち二重化ドライバ S3 は、バスの正常/異常を定期的に監視する。この監視は、二重化ドライバ S3 の設定情報に設定された時間間隔で行われる。

【0060】まず二重化ドライバ S3 は、バスの障害監視を実行する (ステップ 301)。この障害監視、即ち実行バスが正常か否かの監視は、D x 1 への実行バスにコマンドを送信することで行われる。このコマンドには、ディスクデータに影響を及ぼす恐れのないステータス読み出しなどのコマンド、例えば SCSI プロトコルの Mode Sense コマンドを用いる。

【0061】二重化ドライバ S3 は、実行したコマンドが正常に終了したか否かにより、D x 1 への実行バスが正常であるか否かを判定する (ステップ 302)。もし、実行したコマンドが正常終了した場合、二重化ドライバ S3 は実行バスが正常であると判定して処理を完了し、アイドル状態 (ステップ 100) に戻る。

【0062】これに対し、実行したコマンドが異常終了した場合には、二重化ドライバ S3 は D x 1 への実行バスが異常であると判定する。この場合、二重化ドライバ S3 は前記ステップ 203 と同様に、D x 1 へのバスが二重化されているか否かを判定する (ステップ 303)

もし、D x 1 へのバスが二重化されている場合には、二重化ドライバ S3 は前記ステップ 204 と同様に、待機バス (バス E とする) の状態が正常であるか否かを判定する (ステップ 304)。そして待機バスが正常であるならば、二重化ドライバ S3 は前記ステップ 205 と同様に、待機バスと実行バスとを切り替えて、今まで待機バスであったバスを実行バスに、実行バスであったバスを待機バスに変更する (ステップ 305)。

【0063】一方、D x 1 へのバスが二重化されていない場合、或いは二重化されていても待機バスが異常な場合には、二重化ドライバ S3 は障害監視処理にてバスの障害を検出したことをオペレーティングシステム S2 に通知して (ステップ 306)、アイドル状態 (ステップ 100) に戻る。

【0064】以上に述べたように本実施形態においては、ホスト (ホストコンピュータ) C1 x の二重化ドライバ S3 に、従来の SCSI ドライバ S4 経由でのデータ転送用ドライバ S5 の制御に加えて、SCSI ドライバ S4 1 経由での管理用ドライバ S6 の制御を行う機能 (以下、機能 1 と称する) を付加したことを特徴とする。

【0065】次に、ホスト C1 x の管理用ドライバ S6 に、従来の LAN20 を利用しての管理データ (監視データ) 転送に加えて、LAN20 を介して SCSI プロトコルのデータを転送する機能 (以下、機能 2 と称する) を付加したことをも特徴とする。

【0066】そして、ディスクアレイ装置 C2 の管理用ドライバ S13 に、従来の LAN20 を利用しての管理データ (監視データ) 転送に加えて、LAN20 を介して SCSI プロトコルのデータを転送する機能 (以下、機能 3 と称する) を付加したことをも特徴とする。

【0067】このような機能 1 乃至 3 の追加により、データ転送用のバス F に加えて、従来であれば管理データの転送のみに用いられていた管理用のバスのうち、ホスト C1 x の管理用インタフェース P2 からケーブル E1 x、LAN20 及びケーブル E2 を介してディスクアレイ装置 C2 の管理用インタフェース P5 に至るバス E を、バス F と同様のデータ転送用バスとしても割り当てることができる。この結果、従来の高可用性ストレージシステムの二重化構成とは異なって、二重化のためのハードウェア構成、即ちホスト C1 x 及びディスクアレイ装置 C2 内に、データ転送用インタフェース P1 及び P4 に相当するもう 1 つのデータ転送用インタフェースを設けることも、データ転送用インタフェースケーブル F1 x 及び F2 に相当するデータ転送用インタフェースケーブルを設けることも必要なく、ハードウェア構成の削減が図れ、しかも従来の構成と同等のシステムの可用性を実現できる。

【0068】また、従来の二重化構成を適用したシステムに、上記機能 1 乃至 3 を追加して管理用バスをデータ

転送用バスにも利用できるようにした場合には、バスの三重化構成が実現できるため、同じハードウェア構成でありながら、より可用性の高いシステムを実現することができる。

【0069】以上の実施形態では、独立の管理端末C3が設けられている場合について説明したが、これに限るものではない。例えば、ホストC11～C1nのいずれか1つに、ホストC11～C1n及びディスクアレイ装置C2の状態監視を含むシステム管理を行う機構を持たせるようにしてもよい。本発明では、このようなホスト（管理用ホスト）内に設けられるシステム管理機構も、管理端末と定義する。

【0070】なお、本発明は、上記実施形態に限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で種々に変形することが可能である。更に、上記実施形態には種々の段階の発明が含まれており、開示される複数の構成要件における適宜な組み合わせにより種々の発明が抽出され得る。例えば、実施形態に示される全構成要件から幾つかの構成要件が削除されても、発明が解決しようとする課題の欄で述べた課題が解決でき、発明の効果の欄で述べられている効果が得られる場合には、この構成要件が削除された構成が発明として抽出され得る。

【0071】

【発明の効果】以上詳述したように本発明によれば、データ転送用のバス（第2のバス）に加えて、従来であれば管理データの転送のみに用いられていた管理用接続インタフェースを含む管理用のバスのうちの、ホストコンピュータの管理用インタフェース（第2のインタフェース）から当該管理用接続インタフェースを経由してストレージ装置の管理用インタフェース（第4のインタフェース）に至るバス（第1のバス）をデータ転送用バスとしても割り当てることができる。この結果、データ転送用バスを二重化するのに、二重化のためのハードウェア構成を必要とせず、したがってハードウェア構成の削減が図れ、しかも従来の構成と同等のシステムの可用性を

実現できる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る高可用性ストレージシステムの構成を示すブロック図。

【図2】図1中のホストC1x（x=1～n）の主としてソフトウェア構成を示す図。

【図3】図1中のディスクアレイ装置C2内の主としてコントローラP3のソフトウェア構成を示す図。

【図4】図3中のホストC1xの二重化ドライバS3による、バスが二重化されたデバイスの認識処理を説明するためのフローチャート。

【図5】図3中のホストC1xの二重化ドライバS3による、ディスクアクセス処理及びバスの障害監視処理を説明するためのフローチャート。

15 【符号の説明】

C11～C1n, C1x…ホスト（ホストコンピュータ）

C2…ディスクアレイ装置（ストレージ装置）

C3…管理端末

20 P1…データ転送用インタフェース（第1のインタフェース）

P2…管理用インタフェース（第2のインタフェース）

P3…コントローラ

25 P4…データ転送用インタフェース（第3のインタフェース）

P5…管理用インタフェース（第4のインタフェース）

P6…管理用インタフェース

S2, S8…オペレーティングシステム

S3…二重化ドライバ（監視手段、切り替え制御手段）

30 S5, S12…データ転送用ドライバ

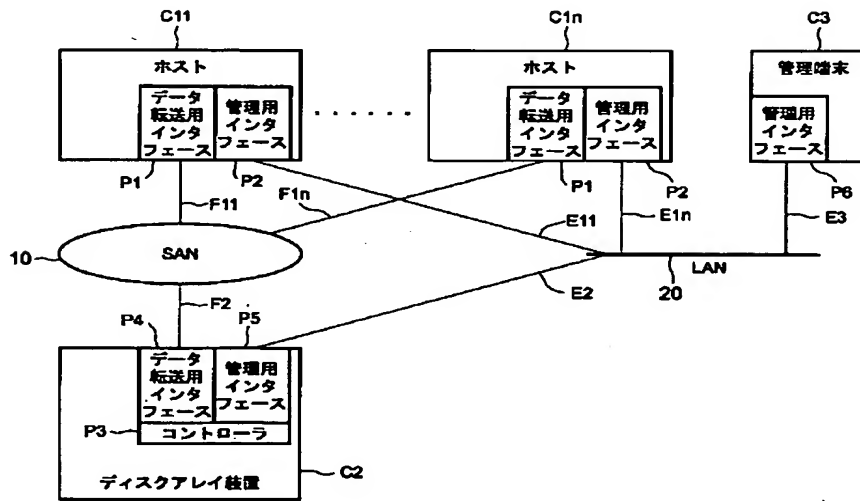
S6…管理用ドライバ（転送制御手段）

S13…管理用ドライバ

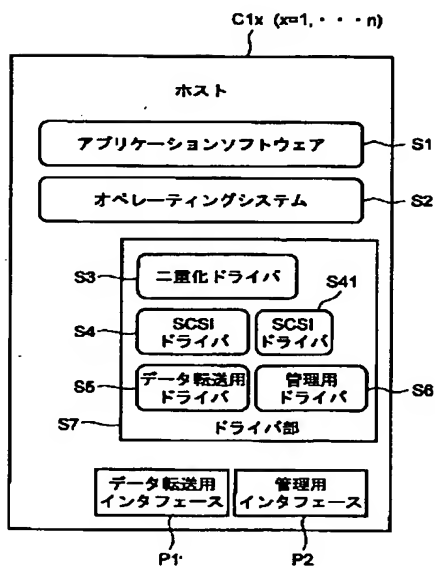
10…SAN（ストレージエリアネットワーク、データ転送用接続インタフェース）

35 20…LAN（ローカルエリアネットワーク、管理用接続インタフェース）

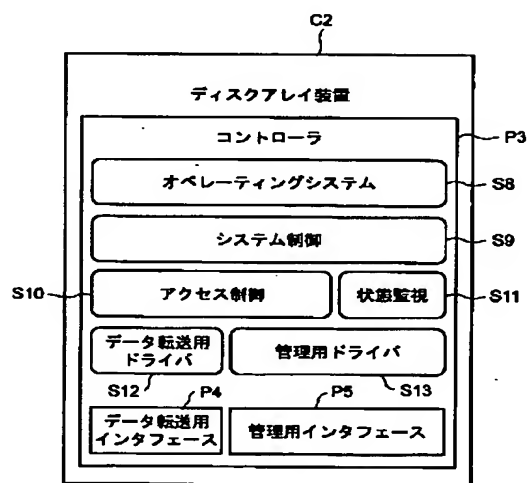
【図1】



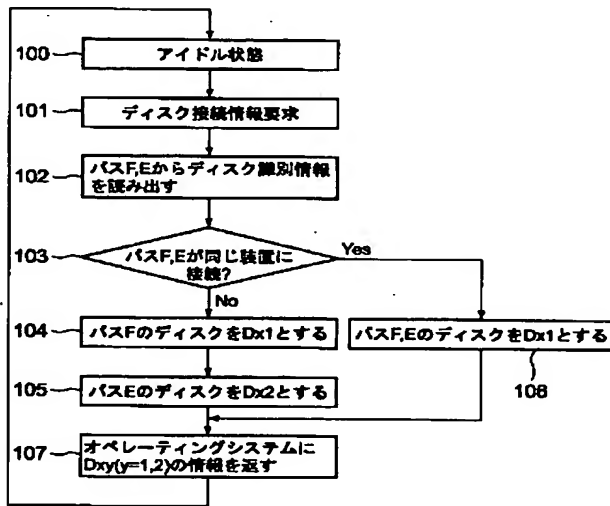
【図2】



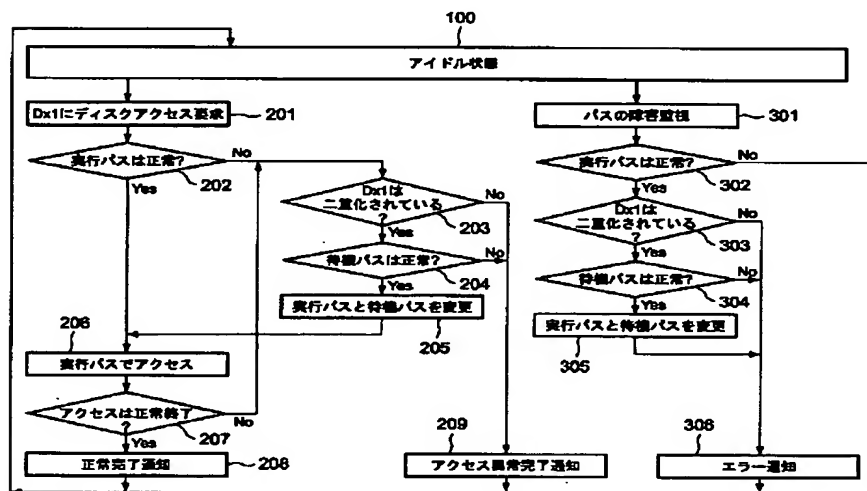
【図3】



【図4】



【図5】



フロントページの続き

(54) 【発明の名称】 管理用インタフェースを用いた高可用性ストレージシステム、同システムに用いられるホストコンピュータ、及び同コンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体